variant sequences. It will generally be preferred that the test mixture contain as large a number of possible sequence variants as is practical for selection, to insure that a maximum number of potential binding sequences are identified. A randomized sequence of 30 nucleotides will contain a calculated 10^{18} different candidate sequences. As a practical matter, it is convenient to sample only about 10^{18} candidates in a single selection. Practical considerations include the number of templates on the DNA synthesis column, and the solubility of RNA and the target in solution. (Of course, there is no theoretical limit for the number of sequences in the candidate mixture.) Therefore, candidate mixtures that have randomized segments longer than 30 contain too many possible sequences for all to be conveniently sampled in one selection. It is not necessary to sample all possible sequences of a candidate mixture to select a nucleic acid ligand of the invention. It is basic to the method that the nucleic acids of the test mixture are capable of being amplified. Thus, it is preferred that any conserved regions employed in the test nucleic acids do not contain sequences which interfere with amplification.

The various RNA motifs described above can almost always be defined by a polynucleotide containing about 30 nucleotides. Because of the physical constraints of the SELEX process, a randomized mixture containing about 30 nucleotides is also about the longest contiguous randomized segment which can be utilized while being able to test substantially all of the potential variants. It is, therefore, a preferred embodiment of this invention when utilizing a candidate mixture with a contiguous randomized region, to use a randomized sequence of at least 15 nucleotides and containing at least about 10^9 nucleic acids, and in the most preferred embodiment contains at least 25 nucleotides.

This invention includes candidate mixtures containing all possible variations of a contiguous randomized segment of at least 15 nucleotides. Each individual member in the candidate mixture may also be comprised of fixed sequences flanking the randomized segment that aid in the amplification of the selected nucleic acid sequences.

Candidate mixtures may also be prepared containing both randomized sequences and fixed sequences wherein the fixed sequences serve a function in addition to the amplification process. In one embodiment of the invention, the fixed sequences in a candidate mixture may be selected in order to enhance the percentage of nucleic acids in the candidate mixture possessing a given nucleic acid motif. For example, the incorporation of the appropriate fixed nucleotides will make it possible to increase the percentage of pseudoknots or hairpin loops in a candidate mixture. A candidate mixture that has been prepared including fixed sequences that enhance the percentage of a given nucleic acid structural motif is, therefore, a part of this invention. One skilled in the art, upon routine inspection of a variety of nucleic antibodies as described herein, will be able to construct, without undue experimentation, such a candidate mixture. Examples 2 and 8 below describe specific examples of candidate mixtures engineered to maximize preferred RNA motifs.

Candidate mixtures containing various fixed sequences or using a purposefully partially randomized sequence may also be employed after a ligand solution or partial ligand solution has been obtained by SELEX. A new SELEX process may then be initiated with a candidate mixture informed by the ligand solution.

Polymerase chain reaction (PCR) is an exemplary method for amplifying of nucleic acids. Descriptions of PCR methods are found, for example in Saiki et al. (1985) Science 230:1350-1354; Saiki et al. (1986) Nature 324:163-166; Scharf et al. (1986) Science 233:1076-1078; Innis et al. (1988) Proc. Natl. Acad. Sci. 85:9436-9440; and in U.S. Pat. No. 4,683,195 (Mullis et al.) and U.S. Pat. No. 4,683,202 (Mullis et al.). IN its basic form, PCR amplification involves repeated cycles of replication of a desired single-stranded DNA (or cDNA copy of an RNA) employing specific oligonucleotide primers complementary to the 3' and 5' ends of the ssDNA, primer extension with a DNA polymerase, and DNA denaturation. Products generated by extension from one primer serve as templates for extension from the other primer. A related amplification method described in PCT published application WO 89/01050 (Burg et al.) requires the presence or introduction of a promoter sequence upstream of the sequence to be amplified, to give a double-stranded intermediate. Multiple RNA copies of the double-stranded promoter containing intermediate are then produced using RNA polymerase. The resultant RNA copies are treated with reverse transcriptase to produce additional double-stranded promoter containing intermediates which can then be subject to another round of amplification with RNA polymerase. Alternative methods of amplification include among others cloning of selected DNAs or cDNA copies of selected RNAs into an appropriate vector and introduction of that vector into a host organism where the vector and the cloned DNAs are replicated and thus amplified (Guatelli, J. C. et al. (1990) proc. Natl. Acad Sci. 87:1874). In general, any means that will allow faithful, efficient amplification of selected nucleic acid sequences can be employed in the method of the present invention. It is only necessary that the proportionate representation of sequences after amplification at least roughly reflects the relative proportions of sequences in the mixture before amplification.

Specific embodiments of the present invention for amplifying RNAs were based on Innis et al. (1988) supra. The RNA molecules and target molecules in the test mixture were designed to provide, after amplification and PCR, essential T7 promoter sequences in their 5' portions. Full-length cDNA copies of selected RNA molecules were made using reverse transcriptase primed with an ologimer complementary to the 3' sequences of the selected RNAs. The resultant cDNAS were amplified by Taq DNA polymerase chain extension, providing the T7 promoter sequences in the selected DNAs. Double-stranded products of this amplification proces were then transcribed in vitro. Transcripts were used in the next selection/amplification cycle. The method can optionally include appropriate nucleic acid purification steps.

In general any protocol which will allow selection of nucleic acids based on their ability to bind specifically to another molecule, i.e., a protein or in the most general case any target molecule, can be employed in the method of the present invention. It is only necessary that the selection partition nucleic acids which are capable of being amplified. For example, a filter binding selection, as described in Example 1, in which a test nucleic acid mixture is incubated with target protein, the nucleic acid/protein mixture is then filtered through

a nitrocellulose filter and washed with appropriate buffer to remove free nucleic acids. Protein/nucleic acid often remain bound to the filter. The relative concentrati ns f protein to test nucleic acid in the incubated mixture influences the strength of binding that is selected for. When nucleic acid is in excess, competition for available binding sites occurs and those nucleic acids which bind most strongly are selected. Conversely, when an excess of protein is employed, it is expected that any nucleic acid that binds to the protein will be selected. The relative concentrations of protein to nucleic acid employed to achieve the desired selection will depend on the type of protein, the strength of the binding interaction and the level of any background binding that is present. The relative concentrations needed to achieve the desired selection result can be readily determined empirically without under experimentation. Similarly, it may be necessary to optimize the filter washing procedure to minimize background binding. Again such optimization of the filter washing procedures is within the skill of the ordinary artisan.

A mathematical evaluation of SELEX referred to as SELEXION has been utilized by the inventors of the present invention. Appendix A to this application includes a brief review of the mathematical analysis utilized to obtain generalizations regarding SELEX derived from SELEXION.

The generalizations obtained from SELEXION are as follows: 1) The likelihood of recovering the best-binding RNA in each round of SELEX increases with the number of such molecules present, with their binding advantage versus the bulk RNA pool, and with the total amount of protein used. Although it is not always intuitively obvious to know in advance how to maximize the difference in binding, the likelihood of recovering the best-binding RNA still can be increased by maximizing the number of RNA molecules and target molecules sampled; 2) the ideal nucleic acid and protein concentrations to be used in various rounds of SELEX are dependent on several factors. The experimental parameters suggested by SELEXION parallel those employed in the Examples hereto. For example, when the relative affinity of the ultimate ligand solution is not known—which will almost inevitably be the case when SELEX is performed—it is preferred that the protein and nucleic acid candidate mixture concentrations are selected to provide a binding between about 3 and 7 percent of the total of nucleic acids to the protein target. By using this criterion it can be expected that a tenfold to twentyfold enrichment in high affinity ligands will be achieved in each round of SELEX.

The experimental conditions used to select nucleic acid ligands to various targets in the preferred embodiment are to be selected to mimic the environment that the target would be found in vivo. Example 10 below indicates how changing the selection conditions will effect the ligand solution received to a particular target. Although the ligand solution to NGF had significant similarities under high and low salt conditions, differences were observed. Adjustable conditions that may be altered to more accurately reflect the in vivo environment of the target include, but are not limited to, the total ionic strength, the concentration of bivalent cations and the pH of the solution. One skilled in the art would be able to easily select the appropriate separation conditions based on a knowledge of the given target.

In order to proceed to the amplification step, selected nucleic acids must be released from the target after partitioning. This process must be done without chemical degradation of the selected nucleic acids and must result in amplifiable nucleic acids. In a specific embodiment, selected RNA molecules were eluted from nitrocellulose filters using a freshly made solution containing 200 μl of a 7M urea, 20 mM sodium citrate (pH 5.0), 1 mM EDTA solution combined with 500 μl of phenol (equilibrated with 0.1M sodium acetate pH 5.2). A solution of 200 μl 7M urea with 500 μl of phenol has been successfully employed. The eluted solution of selected RNA was then extracted with ether, ethanol precipitated and the precipitate was resuspended in water. A number of different buffer conditions for elution of selected RNA from the filters can be used. For example, without limitation nondetergent aqueous protein denaturing agents such as quanidinium chloride, quanidinium thiocyanate, etc., as are known in the art, can be used. The specific solution used for elution of nucleic acids from the filter can be routinely selected by one of ordinary skill in the art.

Alternative partitioning protocols for separating nucleic acids bound to targets, particularly proteins, are available to the art. For example, binding and partitioning can be achieved by passage of the test nucleic acid mixture through a column which contains the target molecule bound to a solid support material. Those nucleic acid that bind to the target will be retained on the column and unbound nucleic acids can be washed from the column.

Throughout this application, the SELEX process has been defined as an iterative process wherein selection and amplification are repeated until a desired selectivity has been attained. In one embodiment of the invention, the selection process may be efficient enough to provide a ligand solution after only one separation step. For example, in theory a column supporting the target through which the candidate mixture is introduced—under the proper conditions and with a long enough column—should be capable of separating nucleic acids based on affinity to the target sufficiently to obtain a ligand solution. To the extent that the original selection step is sufficiently selective to yield a ligand solution after only one step, such a process would also be included within the scope of this invention.

In one embodiment of this invention, SELEX is iteratively performed until a single or a discrete small number of nucleic acid ligands remain in the candidate mixture following amplification. In such cases, the ligand solution will be represented as a single nucleic acid sequence, and will not include a family of sequences having comparable binding affinities to the target.

In an alternate embodiment of the invention, SELEX iterations are terminated at some point when the candidate mixture has been enriched in higher binding affinity nucleic acid ligands, but still contains a relatively large number of distinct sequences. This point can be determined by one of skill in the art by periodically analyzing the sequence randomness of the bulk candidate mixture, or by assaying bulk affinity to the target.

At this time, SELEX is terminated, and clones are prepared and sequenced. Of course, there will be an almost unlimited number of clones that could be sequenced. As seen in the Examples below, however, after sequencing between 20 and 50 clones it is generally possible to detect the most predominant sequences and defining characteristics of the ligand solution. In a hypothetical example, after cloning 30 sequences it will be found that 6 sequences are identical, while certain se-

Secondary selection methods that can be combined with SELEX include among others selections or screens for enzyme inhibition, alteration of substrate binding, loss of functionality, disruption of structure, etc. Those of ordinary skill in the art are able to select among various alternatives those selection or screening methods that are compatible with the methods described herein.

It will be readily apparent to those of skill in the art that in some cases, i.e., for certain target molecules or for certain applications, it may be preferred to employ RNA molecules in preference to DNA molecules as ligands, while in other cases DNA ligands may be preferred to RNA.

The selection methods of the present invention can also be employed to select nucleic acids which bind specifically to a molecular complex, for example to a substrate/protein or inhibitor/protein complex. Among those nucleic acids that bind specifically to the complex molecules, but not the uncompleted molecules there are nucleic acids which will inhibit the formation of the complex. For example, among those nucleic acids ligands which are selected for specific binding to a substrate/enzyme complex there are nucleic acids which can be readily selected which will inhibit substrate binding to the enzyme and thus inhibit or disrupt catalysis by the enzyme.

An embodiment of the present invention, which is particularly useful for the identification or isolation of nucleic acids which bind to a particular functional or active site in a protein, or other target molecule, employs a molecule known, or selected, for binding to a desired site within the target protein to direct the selection/amplification process to a subset of nucleic acid ligands that bind at or near the desired site within the target molecule. In a simple example, a nucleic acid sequence known to bind to a desired site in a target molecule is incorporated near the randomized region of all nucleic acids being tested for binding. SELEX is then used (FIG. 9) to select those variants, all of which will contain the known binding sequence, which bind most strongly to the target molecule. A longer binding sequence, which is anticipated to either bind more strongly to the target molecule or more specifically to the target can thus be selected. The longer binding sequence can then be introduced near the randomized region of the nucleic acid test mixture and the selection/amplification steps repeated to select an even longer binding sequence. Iteration of these steps (i.e., incorporation of selected sequence into test mixtures followed by selection/amplification for improved or more specific binding) can be repeated until a desired level of binding strength or specificity is achieved. This iterative "walking" procedure allows the selection of nucleic acids highly specific for a particular target molecule or site within a target molecule. Another embodiment of such an iterative "walking" procedure, employs an "anchor" molecule which is not necessarily a nucleic acid (see FIGS. 10 and 11). In this embodiment a molecule which binds to a desired target, for example a substrate or inhibitor of a target enzyme, is chemically modified such that it can be covalently linked to an oligonucleotide of known sequence (the "guide oligonucleotide" of FIG. 10). The guide oligonucleotide chemically linked to the "anchor" molecule that binds to the target also binds to the target molecule. The sequence complement of guide oligonucleotide is incorporated near the randomized region of the test nucleic acid mixture. SELEX is then performed to select for those sequences that bind most strongly to the target molecule/anchor complex. The iterative walking procedure can then be employed to select or produce longer and longer nucleic acid molecules with enhanced strength of binding or specifity of binding to the target. The use of the "anchor" procedure is expected to allow more rapid isolation of nucleic acid ligands that bind at or near a desired site within a target molecule. In particular, it is expected that the "anchor" method in combination with iterative "walking" procedures will result in nucleic acids which are highly specific inhibitors of protein function (FIG. 11).

In certain embodiments of the performance of SELEX it is desireable to perform plus/minus screening in conjunction with the selection process to assure that the selection process is not being skewed by some factor unrelated to the affinity of the nucleic acid sequences to the target. For example, when selection is performed by protein binding nitrocellulose, it has been seen that certain nucleic acid sequences are preferentially retained by nitrocellulose and can be selected during the SELEX process. These sequences can be removed from the candidate mixture by incorporating additional steps wherein the preceding SELEX mixture is passed through nitrocellulose to selectively remove those sequences selected solely for that property. Such screening and selection may be performed whenever the target contains impurities or the selection process introduces biases unrelated to affinity to the target.

SELEX has been demonstrated by application to the isolation of RNA molecules which bind to and inhibit the function of bacteriophage T4 DNA polymerase, also termed gp43. The novel RNA ligand of T4 DNA polymerase is useful as a specific assay reagent for T4 DNA polymerase. The synthesis of T4 DNA polymerase is autogenously regulated. In the absence of functional protein, amber fragments and mutant proteins are overexpressed when compared to the rate of synthesis of wild-type protein in replication-deficient infections (Russel (1973) J. Mol. Biol. 79:83-94). In vitro translation of an N-terminal fragment of gp43 is specifically repressed by the addition of purified gp43, and gp43 protects a discrete portion of the mRNA near its ribosome binding site from nuclease attack (Andrake et al. (1988) Proc. Natl. Acad. Sci. USA 85:7942-7946). The size and sequence of the RNA translational operator to which gp43 binds and the strength of that binding have been established. The minimal size of the gp43 operator is a sequence of about 36 nucleotides, as illustrated in FIG. 1, which is predicted to have a hairpin loop structure as indicated therein. The minimal size of the operator was determined by analysis of binding of end-labeled hydrolysis fragments of the operator to gp43. Analysis of binding of operator mutants in the hairpin and loop sequence indicate that gp43 binding to the operator is sensitive to primary base changes in the helix. Binding to the polymerase was even more reduced by changes which significantly reduce hairpin stability. Operator binding was found to be very sensitive to loop sequence. It was found that replication and operator binding in gp43 are mutually exclusive activities. The addition of micromolar amounts of purified RNAs containing intact operator was found to strongly inhibit in vitro replication by gp43.

The wild-type gp43 operator, FIG. 1, was employed as the basis for the design of an initial mixture of RNA molecules containing a randomized sequence region to

assess the ability of the selection/amplification process to isolate nucleic acid molecules that bind to a protein. The RNA test mixture was prepared by in vitro transcription from a 110 base single-stranded DNA template. The template was constructed as illustrated in FIG. 1 to encode most of the wild-type operator sequence, except for the loop sequence. The eight base loop sequence was replaced by a randomized sequence region which was synthesized to be fully random at each base. The template also contained sequences necessary for efficient amplification: a sequence at its 3' end complementarily to a primer for reverse transcription and amplification in polymerase chain reactions and a sequence in its 5' end required for T7 RNA polymerase transcriptional initiation and sufficient sequence complementary to the cDNA of the in vitro transcript. The DNA template is this a mixture of all loop sequence variants, theoretically containing 65,536 individual species.

The dissociation constant for the wild-type loop RNA was found to be about 5×10^{-9}M. The dissociation constant for the population of loop sequence variants was measured to be about 2.5×10^{-7}. Randomization of the loop sequence lowered binding affinity 50-fold.

In vitro transcripts containing the loop sequence variants were mixed with purified gp43 and incubated. The mixture was filtered through a nitrocellulose filter. Protein-RNA complexes are retained on the filter and unbound RNA is not. Selected RNA was then eluted from the filters as described in Example 1. Selected RNAs were extended with AMV reverse transcriptase in the presence of 3' primer as described in Gauss et al. (1987) supra. The resulting cDNA was amplified with Taq DNA polymerase in the presence of the 5' primer for 30 cycles as described in Innis et al. (1986) supra. The selected amplified DNA served as a template for in vitro transcription to produce selected amplified RNA transcripts which were then subject to another round of binding selection/amplification. The RNA/protein ratio in the binding selection mixture was held constant throughout the cycles of selection. The iterative selection/amplification was performed using several different RNA/protein molar ratios. In all experiments RNA was in excess: experiment A employed an RNA/gp43 of 10/1 (moles/moles); experiment B employed an RNA/gp43 of 1000/1; and experiment C employed an RNA/gp43 of 100/1.

The progress of the selection process was monitored by filter binding assays of labelled transcripts of amplified cDNA at the completion of each cycle of the procedure. Batch sequencing of the RNA products from each round for experiment B was also done to monitor the progress of the selection. Autoradiograms of sequencing gels of RNA products after 2, 3 and 4 rounds of selection/amplification are shown in FIG. 3. It is clear that there was no apparent loop sequence bias introduced until after the third selection. After the fourth round of selection, an apparent consensus sequence for the eight base loop sequence is discernable as: A(a/g)(u/c)AAC(u/c)(u/c). Batch sequencing of selected RNA after the fourth round of selection for experiments A, B and C is compared in FIG. 4. All three independent SELEX procedures using different RNA/protein ratios gave similar apparent consensus sequences. There was, however, some apparent bias for wild-type loop sequence (AAUAACUC) in the selected RNA from experiments A and C.

In order to determine what allowable sequence combinations were actually present in the selected RNAs, individual DNAs were cloned from selected RNAs after the fourth round of selection in experiment B. The batch sequence result from experiment B appeared to indicate an even distribution of the two allowable nucleotides which composed each of the four variable positions of the loop sequence. Individuals were cloned into pUC18 as described by Sambrook, J. et al. (1989) Molecular *Cloning: A Laboratory Manual.* (Cold Spring Harbor, N.Y.), Sections 1.13; 1.85-1.86. Twenty individual clones that were identified by colony filter hybridization to the 3' primer were sequenced. None of the sequenced clones were mutant at any place in the operator sequence outside of the loop sequence. Only five variant sequences were observed as shown in FIG. 7, and surprisingly only two sequence variants were the major components of the selected mixture. The frequencies of each sequence in the 20 individual isolates sequenced are also given in FIG. 7. The wild-type sequence AAUAACUC and the loop AGCAACCU were present in approximately equal amount in the selected RNA of experiment B. The other selected variants were 1 base mutants of the two major variants. The strength of binding of the sequence variants was compared in filter binding assays using labelled in vitro transcripts derived from each of the purified clonal isolates. As shown in FIG. 6, a rough correlation between binding affinity of an RNA for gp43 and the abundance of the selected sequence was observed. The two major loop sequence variants showed approximately equal binding affinities for gp43.

The loop sequence variant RNAs isolated by the selection/amplification process, shown in FIG. 7, can all act as inhibitors of gp43 polymerase activity as has been demonstrated for the wild-type operator sequence.

An example of the use of SELEX has been provided by selection of a novel RNA ligand of bacteriophage T4 DNA polymerase (gp43) (Andrake et al. (1988) proc. Natl. Acad. Sci. USA 85:7942-7946).

The present invention includes specific ligand solutions, derived via the SELEX process, that are shown to have an increased affinity to HIV-1 reverse transcriptase, R17 coat protein, HIV-1 rev protein, HSV DNA polymerase, *E.* coli ribosomal protein S1, tPA and NGF. These ligand solutions can be utilized by one of skill in the art to synthesize nucleic acid antibodies to the various targets.

The following examples describe the successful application of SELEX to a wide variety of targets. The targets may generally be divided into two categories—those that are nucleic acid binding proteins and those proteins not known to interact with nucleic acids. In each case a ligand solution is obtained. In some cases it is possible to represent the ligand solution as a nucleic acid motif such as a hairpin loop, an asymmetric bulge or a pseudoknot. In other examples the ligand solution is presented as a primary sequence. In such cases it is not meant to be implied that the ligand solution does not contain a definitive tertiary structure.

In addition to T4 DNA polymerase, targets on which SELEX has been successfully performed include bacteriophage R17 coat protein, HIV reverse transcriptase (HIV-RT), HIV-1 rev protein, HSV DNA polymerase plus or minus cofactor, *E. coli* ribosomal protein S1, tPA and NGF. The following experiments also describe a protocol for testing the bulk binding affinity of a randomized nucleic acid candidate mixture to a variety of

proteins. Example 7 also describes the immobilization of bradykinin and the results of bulk randomized nucleic acid binding studies n bradykinin.

The examples and illustrations herein are not to be taken as limiting in any way. The fundamental insight underlying the present invention is that nucleic acids as chemical compounds can form a virtually limitless variety of sizes, shapes and configurations and are capable of an enormous repertoire of binding and catalytic functions, of which those known to exist in biological systems are merely a glimpse.

EXAMPLES

The following materials and methods were used throughout.

The transcription vector pT7-2 is commercially available (U.S. Biochemical Company, Cleveland, Ohio). Plasmid pUC18 is described by Norrander et al. (1983) Gene 24:15–27 and is also commercially available from new England Biolabs. All manipulations of DNA to create new recombinant plasmids were as described in Maniatis et al. (1982) *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., except as otherwise noted. DNA olignucleotides were synthesized and purified as described in Gauss et al. (1987) Mol. Gen. Genet. 206:24–34.

In vitro transcriptions with T7 RNA polymerase and RNA gel-purification were performed as described in Milligan et al. (1987) Nucl. Acids Res. 15:8783–8798, except that in labeling reactions the concentrations of ATP, CTP, and GTP were 0.5 mM each, and the UTP concentration was 0.05 mM. The UTP was labeled at the alpha position with ^{32}P at a specific activity of approximately 20 Ci/mmol. Crude mRNA preparations from T4 infections, labeling of oligos, and primer extension with AMV reverse transcriptase were all according to Gauss et al. (1987) supra.

Dilutions of labeled, gel-purified RNA and purified gp43 were made in 200 mM potassium acetate, 50 mM Tris-HCl pH 7.7 at 4° C. In nitrocellulose filter binding assays, purified gp43 was serially diluted and 30 μl aliquots of each dilution of protein were added to 30 μl aliquots of diluted, labeled, gel-purified RNA. The RNA dilution (50 μl) was spotted on a fresh nitrocellulose filter, dried and counted to determine input counts per tube. The concentration of protein in the reactions ranged from 10^{-10}M to 10^{-8}M and the concentration of the RNAs in each experiment was approximately 10^{-12}M. After incubation at 4° C. for 30 minutes, each tube was placed at 37° C. for 3 minutes and 50 μl of each sample filtered through pre-wet nitrocellulose filters (Millipore #HAWP 025 00) and washed with 3 ml of 200 mM potassium acetate, 50 mN Tris-HCl pH 7.7. The filters were dried and counted in Ecolume™ scintillation fluid (ICN Biomedicals, Inc.). Controls were done in the absence of gp43, from which the background (always less than about 5% of the input counts) was determined. From each set of measurements the background was subtracted, and the percent of total input counts remaining on he filters calculated. From each set of data points, a best-fit theoretical bimolecular binding curve was generated using a version of a published program (Caceci and Cacheris, 1984 supra) modified to construct a curve described by the equation,

$$\sigma = A[gp43]/(Kd+[43])$$

where σ is the fraction of the total RNA that is bound to the filter, A is the percent of RNA at which binding saturates (approximately 60% for this protein-RNA interaction), [gp43] is the input gp43 concentration, and Kd is the dissociation constant for the bimolecular reaction. This equation is an algebraic rearrangement of equation [1-5] from Bisswanger (1979) *Theorie und Methoden der Enzymkinetik,* Verlag Chemie, Weinheim, FRG, p. 9 with the simplifying assumption that the concentration of the protein far exceeds the concentration of RNA-protein complexes, an assumption which is va id in the experiments described.

EXAMPLE 1. SELECTION OF RNA INHIBITORS OF T4 DNA POLYMERASE

A 110 base single-stranded DNA template for in vitro transcription was created as shown in FIG. 2 by ligation of three synthetic oligonucleotides (Tables 1, 3, 4 and 5) in the presence of two capping oligonucleotides (Tables 1 and 2). One of the template-creating oligos was also used as the 3' primer in reverse transcription of the in vitro transcript and subsequent amplification in polymerase chain reactions (PCRs) (Innis et al. (1988) Proc. Natl. Acad. Sci. USA 85:9436–9440). One of the capping oligos (1) contains the information required for T7 RNA polymerase transcriptional initiation and sufficient sequence complementarity to the cDNA of the in vitro transcript to serve as the 5' primer in the PCR amplification steps. The DNA template encoded an RNA which contains the entire RNA recognition site for T4 DNA polymerase except that a completely random sequence was substituted in place of the sequence which would encode the wild-type loop sequence AAUAACUC. The random sequence was introduced by conventional chemical synthesis using a commercial DNA synthesizer (Applied Biosystems) except that all four dNTP's were present in equimolar amounts in the reaction mixture for each position indicated by N in the sequence of oligonucleotide number 4 (Table 1). The random sequence is flanked by primer annealing sequence information for the 5' and 3' oligos used in PCR. The DNA template is thus a mixture of all loop sequence variants, theoretically containing 65,536 individual species. The dissociation constant for the wild-type loop variant RNA sequence is about 5×10^{-9}M and for the population of sequences was measured to be about 2.5×10^{-7}M, a 50-fold lower binding affinity.

TABLE 1

1) 5'-TAATACGACTCACTATAGGGAGCCAACACCACAATTCCAATCAAG-3'	(SEQ ID NO: 4)
2) 5'-GGGCTATAAACTAAGGAATATCTATGAAAG-3'	(SEQ ID NO: 5)
3) 5'-GAATTGTGGTGTTGGCTCCCTATAGTGAGTCGTATTA-3'	(SEQ ID NO: 6)
4) 5'-ATATTCCTTAGTTTATAGCCCNNNNNNNNAGGCTCTTGATTG-3' and	(SEQ ID NO: 7)
5) 5'-GTTTCAATAGAGATATAAAATTCTTTCATAG-3'	(SEQ ID NO: 8)

In vitro transcripts containing the loop sequence variants were mixed with purified gp43 at three different RNA-protein ratios throughout the multiple rounds of selection. (For A and B the concentration of gp43 was 3×10^{-8}M, "low protein," and for C the concentration of gp43 was 3×10^{-7}M, "high protein." For A the concentration of RNA was about 3×10^{-7}, "low